

机器学习导论

习题五

151250104, 卢以宁, kiwiloveskiwis@gmail.com

2017 年 5 月 17 日

1 [25pts] Bayes Optimal Classifier

试证明在二分类问题中, 但两类数据同先验、满足高斯分布且协方差相等时, LDA 可产生贝叶斯最优分类器。

Solution. 此处用于写证明 (中英文均可)

2 [25pts] Naive Bayes

考虑下面的 400 个训练数据的数据统计情况, 其中特征维度为 2 ($\mathbf{x} = [x_1, x_2]$), 每种特征取值 0 或 1, 类别标记 $y \in \{-1, +1\}$ 。详细信息如表??所示。

根据该数据统计情况, 请分别利用直接查表的方式和朴素贝叶斯分类器给出 $\mathbf{x} = [1, 0]$ 的测试样本的类别预测, 并写出具体的推导过程。

表 1: 数据统计信息

x_1	x_2	$y = +1$	$y = -1$
0	0	90	10
0	1	90	10
1	0	51	49
1	1	40	60

Solution. 此处用于写解答 (中英文均可)

3 [25pts] Bayesian Network

贝叶斯网 (Bayesian Network) 是一种经典的概率图模型，请学习书本 7.5 节内容回答下面的问题：

(1) [5pts] 请画出下面的联合概率分布的分解式对应的贝叶斯网结构：

$$\Pr(A, B, C, D, E, F) = \Pr(A) \Pr(B) \Pr(C) \Pr(D|A) \Pr(E|A) \Pr(F|B, D) \Pr(G|D, E)$$

(2) [5pts] 请写出图??中贝叶斯网结构的联合概率分布的分解表达式。

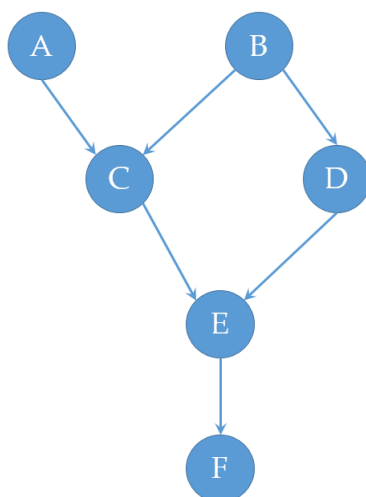


图 1: 题目 3-(2) 有向图

(3) [15pts] 基于第 (2) 问中的图??，请判断表格??中的论断是否正确，只需将下面的表格填完整即可。

表 2: 判断表格中的论断是否正确

序号	关系	True/False	序号	关系	True/False
1	$A \perp\!\!\!\perp B$		7	$F \perp\!\!\!\perp B C$	
2	$A \perp\!\!\!\perp B C$		8	$F \perp\!\!\!\perp B C, D$	
3	$C \perp\!\!\!\perp D$		9	$F \perp\!\!\!\perp B E$	
4	$C \perp\!\!\!\perp D E$		10	$A \perp\!\!\!\perp F$	
5	$C \perp\!\!\!\perp D B, F$		11	$A \perp\!\!\!\perp F C$	
6	$F \perp\!\!\!\perp B$		12	$A \perp\!\!\!\perp F D$	

Solution. 此处用于写解答 (中英文均可)

4 [25pts] Naive Bayes in Practice

请实现朴素贝叶斯分类器，同时支持离散属性和连续属性。详细编程题指南请参见链接：http://lamda.nju.edu.cn/ml2017/PS5/ML5_programming.html.

同时，请简要谈谈你的感想。实践过程中遇到了什么问题，你是如何解决的？

Solution. 此处用于写解答 (中英文均可)